DMB

Display Date 9-23-99
Publication Date 9-24-99
Certifier Jen Wunder

DEPARTMENT OF HEALTH AND HUMAN SERVICES

Food and Drug Administration

[Docket No. 99D–3082]

International Conference on Harmonisation; Choice of Control Group in Clinical Trials

AGENCY: Food and Drug Administration, HHS.

ACTION: Notice.

SUMMARY: The Food and Drug Administration (FDA) is publishing a draft guidance entitled ''E10 Choice of Control Group in Clinical Trials.'' The draft guidance was prepared under the auspices of the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). The draft guidance sets forth general principles that are relevant to all controlled trials and are especially pertinent to the major clinical trials intended to demonstrate drug (including biological drug) efficacy. The draft guidance describes the principal types of control groups and discusses their appropriateness in particular situations. The draft guidance is intended to assist sponsors and investigators in the choice of control groups for clinical trials.

DATES: Written comments by (*insert date 90 days after date of publication in the* **Federal Register**).

ADDRESSES: Submit written comments on the draft guidance to the Dockets Management Branch (HFA–305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. Copies of the draft guidance are available from the Drug Information Branch (HFD–210), Center for Drug Evaluation and Research, Food and Drug Administration, 5600 Fishers Lane, Rockville, MD 20857, 301–827–4573. Single copies of the guidance may be obtained by mail from the Office of Communication, Training and Manufacturers Assistance (HFM–40), Center for Biologics

NAD1

Evaluation and Research (CBER), or by calling the CBER Voice Information System at 1–800–835–4709 or 301–827–1800. Copies may be obtained from CBER's FAX Information System at 1–888–CBER–FAX or 301–827–3844.

FOR FURTHER INFORMATION CONTACT:

Regarding the guidance: Robert Temple, Center for Drug Evaluation and Research (HFD–4), Food and Drug Administration, 5600 Fishers Lane, Rockville, MD 20857, 301–594–6758.

Regarding the ICH: Janet J. Showalter, Office of Health Affairs (HFY–20), Food and Drug Administration, 5600 Fishers Lane, Rockville, MD 20857, 301–827–0864.

SUPPLEMENTARY INFORMATION: In recent years, many important initiatives have been undertaken by regulatory authorities and industry associations to promote international harmonization of regulatory requirements. FDA has participated in many meetings designed to enhance harmonization and is committed to seeking scientifically based harmonized technical procedures for pharmaceutical development. One of the goals of harmonization is to identify and then reduce differences in technical requirements for drug development among regulatory agencies.

ICH was organized to provide an opportunity for tripartite harmonization initiatives to be developed with input from both regulatory and industry representatives. FDA also seeks input from consumer representatives and others. ICH is concerned with harmonization of technical requirements for the registration of pharmaceutical products among three regions: The European Union, Japan, and the United States. The six ICH sponsors are the European Commission, the European Federation of Pharmaceutical Industries Associations, the Japanese Ministry of Health and Welfare, the Japanese Pharmaceutical Manufacturers Association, the Centers for Drug Evaluation and Research and Biologics Evaluation and Research, FDA, and the Pharmaceutical Research and Manufacturers of America. The ICH Secretariat, which coordinates the preparation of documentation, is provided by the International Federation of Pharmaceutical Manufacturers Associations (IFPMA).

The ICH Steering Committee includes representatives from each of the ICH sponsors and the IFPMA, as well as observers from the World Health Organization, the Canadian Health Protection Branch, and the European Free Trade Area.

In May 1998, the ICH Steering Committee agreed that a draft guidance entitled ''E10 Choice of Control Group in Clinical Trials'' should be made available for public comment. The draft guidance is the product of the Efficacy Expert Working Group of the ICH. Comments about this draft will be considered by FDA and the Efficacy Expert Working Group.

In accordance with FDA's good guidance practices (62 FR 8961, February 27, 1997), this document is now being called a guidance, rather than a guideline.

The draft guidance sets forth general principles that are relevant to all controlled trials and are especially pertinent to the major clinical trials intended to demonstrate drug (including biological drug) efficacy. The draft guidance includes a description of the five principal types of controls, a discussion of two important purposes of clinical trials, and an exploration of the critical issue of assay sensitivity, i.e., whether a trial could have detected a difference between treatments when there was a difference, a particularly important issue in noninferiority/equivalence trials. In addition, the draft guidance presents a detailed description of each type of control and considers, for each: (1) Its ability to minimize bias, (2) ethical and practical issues associated with its use, (3) its usefulness and the quality of inference in particular situations, (4) modifications of study design or combinations with other controls that can resolve ethical, practical, or inferential concerns, and (5) its overall advantages and disadvantages.

This draft guidance represents the agency's current thinking on the choice of control group in clinical trials. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. An alternative approach may be used if such approach satisfies the requirements of the applicable statute, regulations, or both.

Interested persons may, on or before (*insert date 90 days after date of publication in the* **Federal Register**), submit to the Dockets Management Branch (address above) written comments

on the draft guidance. Two copies of any comments are to be submitted, except that individuals may submit one copy. Comments are to be identified with the docket number found in brackets in the heading of this document. The draft guidance and received comments may be seen in the office above between 9 a.m. and 4 p.m., Monday through Friday. An electronic version of this guidance is available on the Internet at ''http://www.fda.gov/cder/guidance/index.htm'' or at CBER's World Wide Web site at ''http://www.fda.gov/cber/publications.htm''.

The text of the draft guidance follows:

E10 Choice of Control Group in Clinical Trials[1]

1.0 Introduction

The choice of control group is always a critical decision in designing a clinical trial. That choice affects the inferences that can be drawn from the trial, the degree to which bias in conducting and analyzing the study can be minimized, the types of subjects that can be recruited and the pace of recruitment, the kind of endpoints that can be studied, the public credibility of the results, the acceptability of the results by regulating authorities, and many other features of the study, its conduct, and its interpretation.

1.1 General Scheme and Purpose of Guidance

The general principles considered in this guidance are relevant to all controlled trials. They are of especially critical importance to the major clinical trials carried out during drug development to demonstrate efficacy. This guidance does not address the regulatory requirements in any region, but describes what studies using each design can demonstrate. Although any of the control groups described and discussed below may be useful and acceptable in studies serving as the basis for registration in at least some circumstances, they are not equally appropriate or useful in particular cases. After a brief description of the five principal kinds of controls (see section 1.3), a discussion of two important purposes of clinical

---

[1] This draft guidance represents the agency's current thinking on the choice of control group in clincal trials. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. An altenative approach may be used if such approach satisfies the requirements of the applicable statute, regulations, or both.

trials (see section 1.4), and an exploration of the critical issue of whether a trial could have detected a difference between treatments when there was a difference in noninferiority/equivalence trials (see section 1.5), the guidance will describe each kind of control group in more detail (see section 2.0–2.5.7) and consider, for each:

- Its ability to minimize bias
- Ethical and practical issues associated with its use
- Its usefulness and the quality of inference in particular situations
- Modifications of study design or combinations with other controls that can resolve ethical, practical, or inferential concerns
- Its overall advantages and disadvantages

Several other ICH guidances are particularly relevant to the choice of control group:

- E3: Structure and Content of Clinical Study Reports
- E4: Dose–Response Information to Support Drug Registration
- E6: Good Clinical Practice: Consolidated Guideline
- E8: General Considerations for Clinical Trials
- E9: Statistical Principles for Clinical Trials

In this guidance, the drug terms ''test drug,'' ''study drug,'' and ''investigational drug'' are considered synonymous and are used interchangeably; similarly, ''active control'' and ''positive control,'' ''clinical trial'' and ''clinical study,'' ''control'' and ''control group;'' and ''treatment'' and ''drug'' are essentially equivalent terms.

1.2 Purpose of Control Group

Control groups have one major purpose: to allow discrimination of patient outcomes (changes in symptoms, signs, or other morbidity) caused by the test drug from outcomes caused by other factors, such as the natural progression of the disease, observer or patient expectations, or other treatment. The control group experience tells us what would have happened to patients if they had not received the test treatment (or what would have happened with a different treatment known to be effective).

If the course of a disease were uniform in a given patient population, or predictable from patient characteristics such that outcome could be predicted reliably for any given subject or group of subjects, results of treatment could simply be compared with the known outcome without treatment. For example, one could assume that pain would have persisted for a defined time, blood pressure would not have changed, depression would have lasted for a defined time, tumors would have progressed, the mortality after an acute infarction would have been the same as previously seen. In unusual cases, the course of illness is in fact predictable in a defined population and it may be possible to use a similar group of patients previously studied as a "historical control" (see section 1.3.5). In most situations, however, a concurrent control group is needed because it is not possible to predict outcome with adequate accuracy.

A concurrent control group is one chosen from the same population as the test group and treated in a defined way as part of the same trial that studies the test drug. The test and control groups should be similar with regard to all baseline and on-treatment variables that could influence outcome other than the study treatment. Failure to achieve this similarity can introduce a bias into the study. Bias here (and as used in ICH E9) means the systematic tendency of any aspects of the design, conduct, analysis, and interpretation of the results of clinical trials to make the estimate of a treatment effect deviate from its true value. Randomization and blinding are the two techniques usually used to prevent such bias and to ensure that the test treatment and control groups are similar at the start of the study and are treated similarly in the course of the study (see ICH E9). Whether a trial design includes these features is a critical determinant of its quality and persuasiveness.

1.2.1 Randomization

Assurance that subject populations are similar in test and control groups is best attained by randomly dividing a single sample population into groups that receive the test or control treatments. Randomization avoids systematic differences between groups with respect to variables that could affect outcome. The inability to eliminate systematic differences is the principal problem of studies without a concurrent randomized control (see external control trials, section 1.3.5). Randomization also provides a sound basis for statistical inference.

1.2.2 Blinding

The groups should not only be similar at baseline, but should be treated and observed similarly during the trial, except for receiving the test and control drug. Clinical trials are often "double-blind" (or "double-masked"), meaning that both subjects and investigators (including analysts of data, sponsors, other clinical trial personnel) are unaware of each subject's assigned treatment, to minimize the potential biases resulting from differences in management, treatment, or assessment of patients, or interpretation of results that could arise as a result of subject or investigator knowledge of the assigned treatment. For example:

- Subjects on active drug might report more favorable outcomes because they expect a benefit or might be more likely to stay in a study if they knew they were on active drug.
- Observers might be less likely to identify and report treatment responses in a no-treatment group or might be more sensitive to a favorable outcome or adverse event in patients receiving active drug.
- Knowledge of treatment assignment could affect vigor of attempts to obtain on-study or followup data.
- Knowledge of treatment assignment could affect decisions about whether a subject should remain on treatment or receive concomitant medications or other ancillary therapy.
- Knowledge of treatment assignment could affect decisions as to whether a given subject's results should be included in an analysis.
- Knowledge of treatment assignment could affect choice of statistical analysis.

Double-blinding is intended to ensure that subjective assessments and decisions are not affected by knowledge of treatment assignment.

1.3 Types of Controls

Control groups in clinical trials can be classified on the basis of two critical attributes: (1) The type of treatment received and (2) the method of determining who will be in the control group. The type of treatment may be any of the following four: (1) Placebo, (2) no treatment, (3) different dose or regimen of the study treatment, or (4) different active treatment. The principal methods of determining who will be in the control group are by randomization or by selection of a control population separate from the

population treated in the trial (external or historical control). This document categorizes control groups into five types. The first four are concurrently controlled (the control group and test groups are chosen from the same population and treated concurrently), usually with random assignment to treatment, and are distinguished by which of the types of control treatments listed above are received. External (historical) control groups, regardless of the comparator treatment, are considered together as the fifth type because of serious concerns about the ability to ensure comparability of test and control groups in such trials and the ability to minimize important biases, making this design usable only in exceptional circumstances.

It is increasingly common to carry out studies that have more than one kind of control group. Each kind of control is appropriate in some circumstances, but none is usable or adequate in every situation. The five kinds of control are:

1.3.1 Placebo Concurrent Control

In a placebo-controlled study, subjects are randomly assigned to a test treatment or to an identical-appearing inactive treatment. The treatments may be titrated to effect or tolerance, or may be given at one or more fixed doses. Such trials are almost always double-blind, with both subjects and investigator unaware of treatment assignment. The name of the control suggests that its purpose is to control for "placebo" effect (improvement in a subject resulting from knowing that he or she is taking a drug), but that is not its only or major benefit. Rather, the placebo concurrent control design, by allowing blinding and randomization and including a group that receives no treatment, controls for all potential influences on the actual or apparent course of the disease other than those arising from the pharmacologic action of the test drug. These influences include spontaneous change (natural history of the disease), subject or investigator expectations, use of other therapy, and subjective elements of diagnosis or assessment. Placebo-controlled trials seek to show a difference between treatments when they are studying effectiveness, but may also seek to show lack of difference (of specified size) in evaluating a safety measurement.

1.3.2 No-Treatment Concurrent Control

In a no-treatment controlled study, subjects are randomly assigned to test treatment or to no (i.e., absence of) test or control therapy. The principal difference between this design and a placebo-controlled

trial is that subjects and investigators are not blind to treatment assignment. Because of the advantages of double-blind designs, this design is likely to be needed and suitable only when it is difficult or impossible to double-blind (e.g., medical versus surgical treatment, treatments with easily recognized toxicity) and only when there is reasonable confidence that study endpoints are objective and that the results of the study are unlikely to be influenced by the factors listed in section 1.2.2. Note that it is often possible to blind endpoint assessment, even if the overall trial is not double-blind. This is a valuable approach and should always be considered in studies that cannot be blinded, but it does not solve the other problems associated with knowing the treatment assignment (see section 1.2.2).

1.3.3 Dose-Response Concurrent Control

In a randomized, fixed-dose, dose-response study, subjects are randomized to one of several fixed-dose groups. Subjects may either be placed on their fixed dose initially or be raised to that dose gradually, but the intended comparison is between the groups on their final dose. Dose-response studies are usually double-blind. They may include a placebo (zero dose) and/or active control. In a concentration-controlled trial, treatment groups are titrated to several fixed-concentration windows; this type of trial is conceptually similar to a fixed-dose, dose-response trial.

1.3.4 Active (Positive) Concurrent Control

In an active-control (or positive control) study, subjects are randomly assigned to the test treatment or to an active-control drug. Such trials are usually double-blind, but this is not always possible; many oncology studies, for example, are considered impossible to blind because of different regimens, different routes of administration (see section 1.3.2) and different toxicities. Active-control trials can have two distinct objectives with respect to showing efficacy: (1) To show efficacy of the test drug by showing it is as good as (equivalent, not inferior to) a known effective agent or (2) to show efficacy by showing superiority of the test drug to the active control. They may also be used with the primary objective of comparing the efficacy/safety of the two drugs (see section 1.4). When this design is used to show equivalence/noninferiority or to compare the drugs, it raises the critical question of whether the trial was capable of distinguishing active from inactive treatments (see section 1.5).

1.3.5 External Control (Including Historical Control)

An externally controlled study compares a group of subjects receiving the test treatment with a group of patients external to the study, rather than to an internal control group consisting of patients from the same population assigned to a different treatment. External controls can be a group of patients treated at an earlier time (historical control) or during the same time period but in another setting. The external control may be defined (a specific group of patients) or nondefined (a comparator group based on general medical knowledge of outcome). Use of this latter comparator is particularly treacherous (such trials are sometimes called uncontrolled) because general impressions are so often inaccurate. Baseline-controlled studies, in which subjects' status on therapy is compared with status before therapy (e.g., blood pressure, tumor size), are a variation of this type of control. In this case, the changes from baseline are often compared to a general impression of what would have happened without intervention, rather than to a specific historical experience, although a more defined experience can also be used.

1.3.6 Multiple-Control Groups

As will be described further below (see section 1.5.1), it is often possible and advantageous to use more than one kind of control in a single study, e.g., use of both active drug and placebo. Similarly, trials can use several doses of test drug and several doses of active control, with or without placebo. This design may be useful for active drug comparisons where the relative potency of the two drugs is not well established, or where the purpose of the trial is to establish relative potency.

1.4 Purposes of Clinical Trials

Two purposes of clinical trials should be distinguished: (1) Assessment of the efficacy and/or safety of a treatment and (2) assessment of the relative (comparative) efficacy, safety, benefit/risk relationship or utility of two treatments.

1.4.1 Evidence of Efficacy

In some cases, the purpose of a trial is to demonstrate that a test drug has any clinical effect (or an effect of some specified size). A study using any of the control types may demonstrate efficacy of the test drug by showing that it is superior to the control (placebo, low dose, active drug). An active-

control trial may, in addition, demonstrate efficacy in some cases by showing the new drug to be similar in efficacy to a known effective therapy. The known efficacy of the control is then attributed to the new drug. Clinical studies designed to demonstrate efficacy of a new drug by showing that it is similar in efficacy to a standard agent have been called ''equivalence'' trials. Because in this case the finding of interest is one-sided, these are actually noninferiority trials, attempting to show that the new drug is not less effective than the control by more than a defined amount. As the fundamental assumption of such studies is that showing noninferiority is evidence of efficacy, the decision to utilize this trial design necessitates attention to the question of whether the active control can be relied upon to have an effect in the setting of the trial and whether, as a result, the trial can be relied on not to find a truly inferior drug to be noninferior (see section 1.5).

1.4.2 Comparative Efficacy and Safety

In some cases, the focus of the trial is the comparison with another agent, not the efficacy of the test drug per se. Depending on the therapeutic area, these trials may be seen as providing information needed for relative benefit-risk assessment. The active comparator(s) should be acceptable to the region for which the data are meant. Depending on the situation, it may not be necessary to show equivalence or noninferiority; for example, a less effective drug could have safety advantages and thus be considered useful.

Even though the primary focus of such a trial is the comparison of treatments rather than demonstration of efficacy, the cautions described for conducting and interpreting noninferiority trials need to be taken into account (see section 1.5). The ability of the comparative trial to detect a difference between treatments when one exists needs to be established because a trial incapable of distinguishing between treatments that are in fact different cannot provide useful comparative information.

In addition, for the comparative trial to be informative concerning relative benefit and risk, the trial needs to be fair, i.e., each drug should have an opportunity to perform well. In practice, an active-control equivalence/noninferiority trial offered as evidence of efficacy also almost always should provide a fair comparison with the control, because any doubt as to whether the control in the study had its usual effect

would undermine assurance that the trial had assay sensitivity (see section 1.5). Note that fairness is not an issue when the purpose of the trial is to show efficacy by demonstrating superiority to the control (i.e., the trial will show such efficacy even if the comparator is poorly used; such a trial will not, however, show an advantage over the control).

Among aspects of study design that could unfairly favor one treatment group are choice of dose or patient population and selection and timing of endpoints.

1.4.2.1 *Dose.* In comparing the test drug with an active control for the purpose of assessing relative benefit/risk, it is important to choose an appropriate dose and dose regimen of the control. In examining the results of a comparison of two drugs, it is important to consider whether an apparently less effective control drug has been used at too low a dose or whether the apparently less well tolerated control drug has been used at too high a dose. In some cases, to show superior efficacy or safety convincingly it will be necessary to study several doses of the control and perhaps of the test agent, unless the dose of test agent chosen is superior to any dose (or the only recommended dose) of the control and at least as well tolerated.

1.4.2.2 *Patient population.* Selection of subjects for an active-control trial can affect outcome; the population studied should be carefully considered in evaluating what the trial has shown. For example, if subjects are drawn from a population of nonresponders to the standard agents, there would be a bias in favor of the new agent. The results of such a study could not be generalized to the entire population of previously untreated patients. The result is, however, still good evidence of the efficacy of the new drug. Moreover, a formal study of a new drug in nonresponders to other therapy, in which treatment failures are randomized to either the new or failed therapy (so long as this does not place the patients at risk), can provide an excellent demonstration of the value of the new agent in such nonresponders, a clinically valuable observation (see appendix).

Similarly, it is sometimes possible to identify patient subsets more or less likely to have a favorable response or to have an adverse response to a particular drug. For example, blacks respond poorly to the blood pressure effects of beta blockers and angiotensin-converting enzyme inhibitors, so that a comparison

of a new antihypertensive with these drugs in these patients would tend to show superiority of the new drug. It would not be appropriate to conclude that the new drug is generally superior. Again, however, a planned study in a subgroup, with recognition of its limitations and of what conclusion can properly be drawn, could be informative. See the appendix for a general discussion of "enrichment" study designs, studies that choose a subset of the overall population to increase sensitivity of the study or to answer a specific, but narrow, question.

1.4.2.3 *Selection and timing of endpoints.* When two treatments are used for the same disease or condition, they may differentially affect various outcomes of interest in that disease, particularly if they represent different classes or modalities of therapy. Therefore, when comparing them in a clinical trial, the choice and timing of endpoints may favor one therapy or the other. For example, thrombolytics in patients with acute myocardial infarction can reduce mortality but increase stroke risk. If a new, more active thrombolytic were compared with an older thrombolytic, the more active drug might look better if the endpoint were mortality, but worse if the endpoint were a composite of mortality and disabling stroke. Similarly, in comparing two analgesics in the management of dental pain, assigning a particularly heavy weight to pain at early time points would favor the agent with more rapid onset over an agent that provides greater or longer lasting relief.

1.5 Sensitivity-to-Drug-Effects and Assay Sensitivity of Studies Intended to Show Noninferiority/ Equivalence

As noted in section 1.4.1, use of an active-control noninferiority/equivalence design to demonstrate efficacy poses a particular problem, one not found in trials intended to show a difference between treatments. A demonstration of efficacy by showing noninferiority/equivalence of the new therapy to the established effective treatment or, more accurately, by showing that the difference between them is no larger than a specified size (margin), rests on a critical assumption: that if there is a true difference between the treatments, i.e., if the new drug has a much smaller effect or no effect, the study would not have concluded there was no such difference. This assumption, in turn, rests on the assumption that the active-control drug will have had an effect of a defined size in the study. If these assumptions are incorrect,

an erroneous conclusion that a drug is effective may be reached because a trial seeming to support noninferiority will not in fact have done so.

The ability of a specific trial to detect differences between treatments if they exist has been called, and is here termed, "assay sensitivity." In the noninferiority trial setting, assay sensitivity requires that there be an effect of the control drug in the trial of at least a specified size and that, because of the presence of that effect, the trial has an ability not to declare noninferiority of a new drug when the new drug is in fact inferior. As noted, because the actual effect size of the control in the trial is not measured, the presence of assay sensitivity must be deduced. In this document, the term assay sensitivity, a property of a particular trial, is distinguished from sensitivity-to-drug-effects. Sensitivity-to-drug-effects is defined as the ability of appropriately designed and conducted trials in a specific therapeutic area, using a specific active drug (or other drugs with similar effects), to reliably show a drug effect of at least a minimum size under the conditions of the trial. Sensitivity-to-drug-effects is determined from historical experience; it will usually be established by a determination that such trials, when adequately powered, regularly distinguish active drugs from placebo. Sensitivity-to-drug-effects, established in this way, will imply that, in a similarly well-designed and conducted noninferiority trial, there will be an ability not to find an ineffective agent to be noninferior. Assay sensitivity, in contrast, applies to a specific trial and requires the actual presence of a control drug effect and thus the actual ability of the trial not to declare an inferior drug noninferior. This ability depends on the details of the design and conduct of a specific trial, as well as the presence of sensitivity-to-drug-effects.

1.5.1 Need to Ensure Assay Sensitivity in Noninferiority (Equivalence) Trials; Difference–Showing Versus Noninferiority Studies

When designing a noninferiority study, study designers need to consider the fundamental distinction between two kinds of clinical trials: (1) Those that seek to demonstrate efficacy by showing superiority of a treatment to a control (superiority trials) and (2) those that seek to show efficacy by demonstrating that a new treatment is as good as (not inferior by some specified amount to) a treatment known to be effective. In the difference-showing trial, the finding of a difference itself documents the assay sensitivity

of the trial and documents the efficacy of the superior treatment, so long as the inferior treatment, if an active drug, is known to be no worse than a placebo. In the noninferiority situation, in contrast, a finding of noninferiority leaves unanswered the question: Would the study have led to a conclusion of noninferiority even if the study drug were inferior? In a noninferiority trial without a placebo group, there is no internal standard (that is, a showing of an active drug-placebo difference) to measure/ensure assay sensitivity. The existence of assay sensitivity of the trial therefore needs to be deduced or assumed based on past experience (''historically'') with the control drug, generally from placebo-controlled trials, establishing the sensitivity-to-drug-effects of well-designed and conducted trials, together with evidence that the trial was in fact well conducted.

The question of assay sensitivity, although particularly critical in noninferiority studies, actually arises in any trial that fails to detect a difference between treatments, including a placebo-controlled trial. If a drug fails to show superiority to placebo, for example, it means either that the drug was ineffective or that the study was not capable of detecting the effect of the drug. A straightforward solution to the problem of assay sensitivity is the three-arm study, including both placebo and a known active treatment, a study design with several advantages. Such a study measures effect size (test drug versus placebo) and allows comparison of test drug and active control in a setting where assay sensitivity is established by the active control-placebo comparison. The design is also particularly informative when the test drug and placebo give similar results in the study. In that case, if the active control is superior to placebo, the study did have assay sensitivity and the study provides some evidence that the test drug has little or no efficacy. On the other hand, if neither drug, including the known effective active control, can be distinguished from placebo with respect to efficacy, the clinical study lacks assay sensitivity and does not provide evidence that the drug is ineffective.

1.5.2 Choosing the Noninferiority Margin

As noted earlier, most active-control ''equivalence'' trials are really noninferiority trials intended to establish the efficacy of a new drug. Analysis of the results of noninferiority trials is discussed in the ICH guidances E9 and E3. Briefly, in such a trial, new and established therapies are compared. Prior

to the trial, an equivalence or noninferiority margin, sometimes called a "delta," is selected. This margin is the degree of inferiority of the test drug compared to the control that the trial will attempt to exclude statistically. If the confidence interval for the difference between the test and control treatments excludes a degree of inferiority of the test drug as large as, or larger than, the margin, the test drug can be declared noninferior and thus effective; if the confidence interval includes a difference as large as the margin, the test drug cannot be declared noninferior and cannot be considered effective.

The margin chosen for a noninferiority trial cannot be greater than the smallest effect size that the active drug would be reliably expected to have compared with placebo in the setting of the planned trial, but may be smaller based on clinical judgment. If a difference between active control and new drug favors the control by as much as or more than that amount, the new drug might have no effect at all. The margin generally is identified based on past experience in placebo-controlled trials of adequate design under conditions similar to those planned for the new trial. Note that exactly how to calculate the margin is not described in this document, and there is little published experience on how to do this. The determination of the margin is based on both statistical reasoning and clinical judgment, should reflect uncertainties in the evidence on which the choice is based, and should be suitably conservative. If this is done properly, a finding that the confidence interval for the difference between new drug and the active control excludes a suitably chosen margin could provide assurance that the drug has an effect greater than zero. In practice, the margin chosen usually will be smaller than that suggested by the smallest expected effect size of the active control because of interest in ensuring that some particular clinically acceptable effect size (or fraction of the control drug effect) was maintained. This would also be true in a trial whose primary focus is the therapeutic equivalence of a test drug and active control (see section 1.4.2), where it would be usual to seek assurance that the test and control drug were quite similar, not simply that the new drug had any effect at all.

The fact that the choice of the margin to be excluded can only be based on past experience gives the noninferiority trial an element in common with a historically controlled (externally controlled) study. This study design is appropriate and reliable only when the historical estimate of an expected drug effect

can be well supported by reference to the results of previous studies of the control drug. These studies should lead to the conclusion that the active control can consistently be distinguished from placebo in trials of design similar to the proposed trial (patient population, study size, study endpoints, dose, concomitant therapy, etc.) and should identify an effect size that represents the smallest effect that the control can reliably be expected to have. If placebo-controlled trials of a design similar to the one proposed more than occasionally show no difference between the proposed active control and placebo, and this cannot be explained by some characteristic of the study, only superiority of the test drug would be interpretable. Note that it is the estimated difference from placebo, not the total change from baseline, that needs to be used to calculate the expected effect of the control.

1.5.3 Sensitivity-to-Drug-Effects Is Difficult to Support in Many Situations

Whether the historically based assurance of sensitivity-to-drug-effects of a trial is supported in any given case is to some degree a matter of judgment. There are many conditions, however, in which drugs considered effective cannot regularly be shown superior to placebo in well-controlled studies, and one therefore cannot reliably determine a minimum effect the drug will have in the setting of a specific trial. Such conditions tend to include those in which there is substantial improvement and variability in placebo groups, and/or in which the effects of therapy are small, or variable, such as depression, anxiety, dementia, angina, symptomatic congestive heart failure, seasonal allergies, and symptomatic gastroesophageal reflux disease.

In all these cases, there is no doubt that the standard treatments are effective because there are many well-controlled studies of each of these drugs that have shown an effect. Based on available experience, however, it would be difficult to describe study conditions in which the drug would reliably have at least a minimum effect (i.e., conditions in which there is sensitivity-to-drug-effects) and that, therefore, could be used to identify an appropriate margin. In some cases, the experience on which the expectation of sensitivity-to-drug-effects is based may be of questionable relevance, e.g., if standards of treatment and diagnosis have changed substantially over time. If someone proposing to use an active-control noninferiority design cannot provide acceptable support for the sensitivity-to-drug-effects of the study with the chosen

inferiority margin, a finding of noninferiority cannot be considered informative with respect to efficacy or to a showing of clinical comparability/equivalence.

1.5.4 Assay Sensitivity and Study Quality in Noninferiority Designs

Even where historical experience indicates that studies in a particular therapeutic area are likely to have sensitivity-to-drug-effects, this likelihood can be undermined by the particular circumstances under which the study was conducted. Great attention therefore needs to be paid to how the trial was designed and conducted to determine whether it actually did have assay sensitivity. There are many factors that can reduce a trial's assay sensitivity, such as:

1. Poor compliance with therapy

2. Poor responsiveness of the study population to drug effects

3. Use of concomitant medication or other treatment that interferes with the test drug or that reduces the extent of the potential response

4. A population that tends to improve spontaneously, leaving no room for further drug-induced improvement

5. Poor diagnostic criteria (patients lacking the disease to be studied)

6. Inappropriate (insensitive) measures of drug effect

7. Excessive variability of measurements

8. Biased assessment of endpoint because of knowledge that all patients are receiving a potentially active drug, e.g., a tendency to read blood pressure responses as greater than they actually are, reducing the difference between test drug and control

Clinical researchers and trial sponsors intend to perform high quality studies, and the publication of the Good Clinical Practices guidance will enhance study quality. Nonetheless, it should be appreciated that in trials intended to show a difference between treatments there is a strong imperative to utilize a good study design and minimize study errors, because trial imperfections increase the likelihood of failing to show a difference between treatments when one exists. In placebo-controlled trials, for example, there is often a withdrawal period to be sure study subjects actually have the disease for which treatment is

intended, and great care is taken in defining entry criteria to be sure patients have an appropriate stage of the disease. It is common to have a single-blind placebo run-in period to discover and eliminate subjects who recover spontaneously, whose measurements are too variable, or who are likely to comply poorly with the protocol. There is close attention to trial conduct, including administration of the correct treatments to patients, encouraging compliance with medication use, controlling (or at least recording) concomitant drug use and other concomitant illness, and use of standard procedures for measurement (technique, timing, training periods). All of these efforts will help ensure that an effective drug will be distinguished from placebo. Nonetheless, in many clinical settings, despite the strong stimulus and extensive efforts to ensure study excellence and assay sensitivity, clinical studies are often unable to reliably distinguish effective drugs from placebo.

In contrast, in trials intended to show that there is not a difference of a particular size (noninferiority) between two treatments, there is a much weaker stimulus to engage in many of these efforts, which help ensure that differences will be detected, i.e., ensure sensitivity, because failure to show a difference greater than the margin is the desired outcome of the study. Although some kinds of study error diminish observed differences between treatments, it is noted that some kinds of study errors can increase variance, which would decrease the likelihood of showing noninferiority by widening the confidence interval so that a test drug control difference greater than the margin cannot be excluded. There would therefore be a strong stimulus in these trials to reduce variance, which might be caused, for example, by poor measurement technique. Many errors of the kind described, however, reduce the observed difference between treatments (and thus assay) without necessarily increasing variance. They therefore increase the likelihood that an inferior drug will be found noninferior.

When a noninferiority study is offered as evidence of effectiveness of a new drug, both the sponsor and regulatory authority need to pay particularly close attention to study quality. Whether a given study has assay sensitivity often cannot be determined, but the known reasons for failure to have such sensitivity should be monitored. The design and conduct of the study need to be shown to be similar to studies of the active control that were successful in the past. To ensure that sensitivity-to-drug-effects seen in

past studies is likely to be present in the new study, there should be close attention to critical design characteristics such as the entry criteria and characteristics of the study population (severity of medical condition, method of diagnosis), the specific endpoint measured and timing of assessments, and the use of washout periods to exclude patients without disease or to exclude patients with spontaneous improvement. Similarly, aspects of study conduct that could decrease assay sensitivity should also be examined, including such characteristics as compliance with therapy, monitoring of concomitant therapy, enforcement of entry criteria, and prevention of study dropouts.

One other possibility should be considered. Even where a study seems likely to have sensitivity-to-drug-effects based on prior studies, the population studied or other aspects of study design or conduct in a noninferiority study may be so different that results with the active-control treatment are visibly atypical (e.g., cure rate in an antibiotic trial that is unusually high or low). In that case, the results of a noninferiority trial may not be persuasive.

2.0 Detailed Consideration of Types of Control

2.1 Placebo Control

2.1.1 Description (See Section 1.3.1)

In a placebo-controlled study, subjects are assigned, almost always by randomization, to either a test drug or to a placebo. A placebo is a "dummy" medication that appears as identical as possible to the investigational or test drug with respect to physical characteristics such as color, weight, taste and smell, but that does not contain the test drug. Some trials may study more than one dose of the test drug or include both an active control and placebo. In these cases, it may be easier for the investigator to use more than one placebo ("double-dummy") than to try to make all treatments look the same. The use of placebo facilitates, and is almost always accompanied by, double-blinding (or double-masking). The difference in measured outcome between the active drug and placebo groups is the measure of drug effect under the conditions of the study. Within this general description there is a wide variety of designs that can be used successfully: Parallel or cross-over designs (see ICH E9), single fixed dose or titration in the active drug group, several fixed doses. Several designs meriting special attention will be described

below. Note that not every study that includes a placebo is a placebo-controlled study. For example, an active-control study could use a placebo for each drug (double-dummy) to facilitate blinding; this is still an active-control trial, not a placebo-controlled trial. A placebo-controlled trial is one in which treatment with a placebo is compared with treatment with an active drug.

2.1.2 Ability to Minimize Bias

The placebo-controlled trial, using randomization and blinding, generally reduces subject and investigator bias maximally, but such trials are not impervious to blind-breaking through recognition of pharmacologic effects of one treatment (perhaps a greater concern in cross-over designs); blinded outcome assessment can enhance bias reduction in such cases.

2.1.3 Ethical Issues

When a new agent is tested for a condition for which no effective treatment is known, there is usually no ethical problem with a study comparing the new agent to placebo. Use of a placebo control may raise problems of ethics, acceptability, and feasibility, however, when an effective treatment is available for the condition under study in a proposed trial. In cases where an available treatment is known to prevent serious harm, such as death or irreversible morbidity in the study population, it is generally inappropriate to use a placebo control. There are occasional exceptions, however, such as cases in which standard therapy has toxicity so severe that many patients will refuse therapy.

In other situations, when there is no major health risk associated with withholding or delay of effective therapy, it is considered ethical to ask patients to participate in a placebo-controlled trial, even if they may experience discomfort as a result, provided the setting is noncoercive and they are fully informed about available therapies and the consequences of delaying treatment. Such trials, however, may pose important practical problems. For example, deferred treatment of pain or other symptoms may be unacceptable to patients or physicians and they may not want to participate in such a study. Whether a particular placebo-controlled trial of a new agent will be acceptable to subjects and investigators when there is known effective therapy is a matter of investigator, patient, and institutional review board (IRB)/ independent ethics committee (IEC) judgment, and acceptability may differ among ICH regions.

Acceptability could depend on the specific design of the study and the patient population chosen, as will be discussed below (see section 2.1.5).

Whether a particular placebo-controlled trial is ethical may, in some cases, depend on what is believed to have been clinically demonstrated and on the particular circumstances of the trial. For example, a short term placebo-controlled study of a new antihypertensive agent in patients with mild essential hypertension and no end-organ disease might be considered generally acceptable, while a longer study, or one that included sicker patients, probably would not be.

It should be noted that use of a placebo or no-treatment control does not imply that the patient does not get any treatment at all. For instance, in an oncology trial, when no active drug is approved, patients in both the placebo/no-treatment group and the test drug group will receive needed palliative treatment, such as analgesics.

2.1.4 Usefulness of Placebo–Controlled Trials and Quality/Validity of Inference in Particular Situations

When used to show effectiveness of a treatment, the placebo-controlled trial is as free of assumptions and need for external (extra-study) information as it is possible to be. Most trial design problems and careless errors result in failure to demonstrate a treatment difference (and thereby establish efficacy), so that the trial contains built-in incentives for study excellence. Even when the primary purpose of a trial is comparison of two active agents or assessment of dose-response, the addition of a placebo provides an internal standard that enhances the inferences that can be drawn from the other comparisons.

Placebo-controlled trials also provide the maximum ability to distinguish adverse effects due to drug from those due to underlying disease or intercurrent illness. Note that where they are used to show similarity, for example, to show the absence of an adverse effect, placebo-controlled trials have the same assay sensitivity problem as any equivalence or noninferiority trial (see section 1.5.1). To interpret the result, one must know that if the study drug caused an adverse event, it would have been observed.

2.1.5 Modifications of Design and Combinations With Other Controls That Can Resolve Ethical, Practical, or Inferential Issues

It is often possible to address the ethical or practical limitations of placebo-controlled trials by using modified study designs that still retain the inferential advantages of these trials. In addition, placebo-controlled trials can be made more informative by inclusion of additional treatment groups, such as multiple doses of the test agent or a known active-control treatment.

2.1.5.1 *Additional control groups.*

2.1.5.1.1 *Three-arm study; placebo and active control.* As noted in section 1.5.1, three-arm studies including an active-control as well as a placebo-control group can readily assess whether a failure to distinguish test drug from placebo implies ineffectiveness of the test drug or simply a study that lacked the ability to identify an active drug. The placebo-standard drug comparison in such a trial provides internal evidence of assay sensitivity. It is possible to make the active groups larger than the placebo group in order to improve the precision of the active drug comparison, if this is considered important. This may also make the study more appealing to patients, as there is less chance of being randomized to placebo.

2.1.5.1.2 *Additional doses.* Randomization among several fixed doses of the test drug in addition to placebo allows assessment of dose-response and may be particularly useful in a comparative trial to ensure a fair comparison of treatments (see ICH E4: Dose–Response Information to Support Drug Registration).

2.1.5.1.3 *Factorial/combination studies.* Factorial/ combination (response-surface) designs may be used to explore several doses of the investigational drug as monotherapy and in combination with several doses of another agent proposed for use in combination with it. A single study of this type can define the properties of a wide array of combinations. Such studies are common in the evaluation of new antihypertensive therapies, but can be considered in a variety of settings where more than one treatment is used simultaneously. For example, the independent additive effects of aspirin and streptokinase in preventing mortality after a heart attack were shown in such a trial.

2.1.5.2 *Changes in study design.*

2.1.5.2.1 *Add-on study, placebo-controlled; replacement study.* An "add-on" study is a placebo-controlled trial of a new agent conducted in people also receiving standard therapy. Such studies are useful when standard therapy is known to decrease mortality or irreversible morbidity, so that the therapy cannot be withheld from a patient population known to benefit from it, and when a noninferiority trial with standard treatment as the active control cannot be carried out or would be difficult to interpret (see section 1.5). It is common to study anticancer, antiepileptic, and anti-heart-failure drugs this way. This design is useful only when standard therapy is not fully effective (which, however, is almost always the case), and it has the advantage of providing evidence of improved clinical outcomes (rather than "mere" noninferiority). Efficacy is, of course, established by such studies only for combination therapy, and the dose in a monotherapy situation might be different from the dose found to be effective in combination. In general, this approach is likely to succeed only when the new and standard therapies utilize different pharmacologic mechanisms, although there are exceptions. For example, AIDS combination therapies may show a beneficial effect of pharmacologically-related drugs because of delays in development of resistance.

A variation of this design that can sometimes give information on monotherapy and that is particularly applicable in the setting of chronic disease, is the replacement study, in which the new drug or placebo is added by random assignment to conventional treatment given at an effective dose and the conventional treatment is then withdrawn, usually by tapering. The ability to maintain the subjects' baseline status is then observed in the drug and placebo groups using predefined success criteria. This approach has been used to study steroid-sparing substitutions in steroid-dependent patients without need for initial steroid withdrawal and recrudescence of symptoms in a wash-out period, and has also been used to study antiepileptic drug monotherapy.

2.1.5.2.2 *"Early escape"; rescue medication.* It is possible to design a study to plan for "early escape" from ineffective therapy. Early escape refers to prompt removal of subjects whose clinical status worsens or fails to improve to a defined level (blood pressure not controlled by a prespecified time, seizure rate greater than some prescribed value, blood pressure rising to a certain level, angina frequency above a defined level, liver enzymes failing to normalize by a preset time in patients with hepatitis), who have

a single event that treatment was intended to prevent (first recurrence of unstable angina, grand mal seizure, paroxysmal supraventricular arrhythmia), or who otherwise require added therapy. In such cases, the need to change therapy becomes a study endpoint. The criteria for deciding whether these endpoints have occurred should be well specified, and the timing of measurements should ensure that patients will not remain untreated with an active drug while their disease is poorly controlled. The primary difficulty with this trial design is that it may give information only on short-term effectiveness. The randomized withdrawal trial (see section 2.1.5.2.4), however, which can also incorporate early-escape features, can give information on long-term effectiveness. It should be noted that formal use of rescue medication in response to clinical deterioration could be utilized similarly.

2.1.5.2.3 *Limited placebo period.* In a longer term active-control trial, the addition of a placebo group treated for a short period may establish assay sensitivity (at least for short-term effects). The trial would then continue without the placebo group.

2.1.5.2.4 *Randomized withdrawal.* In a randomized withdrawal study, subjects receiving an investigational therapy for a specified time are randomly assigned to continued treatment with the investigational therapy or to placebo (i.e., withdrawal of active therapy). Subjects for such a trial could be derived from an organized open single-arm study, from an existing clinical cohort (but usually with a formal "wash-in" phase to establish the initial on-therapy baseline), from the active arm of a controlled trial, or from one or both arms of an active-control trial. Any difference that emerges between groups receiving continued treatment and placebo would demonstrate the effect of the active treatment. The prerandomization observation period on drug can be of any length; this approach can therefore be used to study long-term persistence of effectiveness when long-term placebo treatment would not be acceptable. The postwithdrawal observation period could be of fixed duration or could use early escape or time to event (e.g., relapse of depression) approaches. As with the early-escape design, procedures for monitoring patients and assessing study endpoints need careful attention to ensure that patients failing on an assigned treatment are identified rapidly.

The randomized withdrawal approach is suitable in several situations. First, it may be suitable for drugs that appear to resolve an episode of recurring illness (e.g., antidepressants), in which case the withdrawal study is in effect a relapse-prevention study. Second, it may be used for drugs that suppress a symptom or sign (chronic pain, hypertension, angina), but where a long-term placebo-controlled trial would be difficult; in this case, the study can establish long-term efficacy. Third, the design can be used to determine how long a therapy should be continued (e.g., postinfarction treatments with a beta-blocker).

The general advantage of randomized withdrawal designs, when used with an early-escape endpoint, such as return of symptoms, is that the period of placebo exposure with poor response that a patient would have to undergo is short.

Dosing issues can be addressed by this type of design. After all patients had received an initial fixed dose, they could be randomly assigned in the ''withdrawal'' phase to several different doses (as well as placebo), a particularly useful approach when there is reason to think the initial and maintenance doses might be different, either on pharmacodynamic grounds or because there is substantial accumulation of active drug resulting from a long half life of parent drug or active metabolite. Note that the randomized withdrawal design could be used to assess dose-response after an initial placebo-controlled titration study. The titration study is an efficient design for establishing effectiveness, but does not give good dose-response information. The randomized withdrawal phase, with responders randomly assigned to several fixed doses and placebo, will study dose-response rigorously while allowing the efficiency of the titration design.

In utilizing randomized withdrawal designs, it is important to appreciate the possibility of withdrawal phenomena, suggesting the wisdom of relatively slow tapering. A patient may develop tolerance to a drug such that no benefit is being accrued, but the drug's withdrawal may lead to disease exacerbation, resulting in an erroneous conclusion of persisting efficacy. It is also important to realize that treatment effects observed in these studies may be larger than those seen in the general population because randomized withdrawal studies are ''enriched'' with responders (see appendix). This phenomenon results when the study explicitly includes only subjects who appear to have responded to the drug or includes only people who have completed a previous phase of study (which is often an indicator of a good response).

2.1.5.2.5 *Other design considerations.* In any placebo-controlled study, unbalanced randomization (e.g., 2:1, study drug to placebo) may enhance the safety data base and may also make the study more attractive to patients and/or investigators.

2.1.6 Advantages of Placebo-Controlled Trials

2.1.6.1 *Ability to demonstrate efficacy credibly.* Like other difference-showing trials, the interpretation of the placebo-controlled study relies on no externally based assumptions of sensitivity-to-drug-effects nor an assessment of assay sensitivity. These may be the only credible study designs in situations where it is not possible to conclude that noninferiority studies would have assay sensitivity (see section 1.5).

2.1.6.2 *Measures "absolute" effectiveness and safety.* The placebo-controlled trial measures the absolute effect of treatment and allows a distinction between adverse events due to the drug and those due to the underlying disease or "background noise." The absolute effect size information is valuable in a three-group trial (test, placebo, active), even if the primary purpose of the trial is the test versus active control comparison.

2.1.6.3 *Efficiency.* Placebo-controlled trials are efficient in that they can detect treatment effects with a smaller sample size than any other type of concurrently controlled study. Active-control trials intended to show superiority of the new treatment are generally seeking smaller differences than the active-placebo difference sought in a placebo-controlled trial, resulting in need for a larger sample size. Noninferiority active-control trials also need larger sample sizes because they must use conservative assumptions about the effect size of the control drug to ensure that noninferiority of the test drug would in fact demonstrate efficacy. Designers of dose-response studies need to guess at the shape and position of the dose-response curve and may wastefully assign some subjects to several doses that have no effect or are on a response plateau.

2.1.6.4 *Minimizing the effect of subject and investigator expectations.* Use of a blinded placebo control may decrease the amount of improvement resulting from subject or investigator expectations because both are aware that some subjects will receive no active drug. This may increase the ability of the study to detect true drug effects.

2.1.7 Disadvantages of Placebo-Controlled Trials

2.1.7.1 *Ethical concerns (see sections 2.1.3 and 2.1.4).* When effective therapy that is known to prevent harm exists for a particular population, that population cannot usually be ethically studied in placebo-controlled trials; the particular conditions and populations for which this is true may be controversial. Ethical concerns may also direct studies toward less ill subjects or cause studies to examine short-term endpoints when long-term outcomes are of greater interest. Where a placebo-controlled trial is unethical and an active-control trial would not be credible, it may be very difficult to study new drugs at all. For example, it would not be considered ethical to carry out a placebo-controlled trial of a beta blocker in postinfarction patients; yet it would be difficult to conclude that a noninferiority trial would have sensitivity-to-drug-effects. The designs described in section 2.1.5 may be useful in some of these cases.

2.1.7.2 *Patient and physician practical concerns.* Physicians and/or patients may be reluctant to accept the possibility that the patient will be assigned to the placebo treatment, even if there is general agreement that withholding or delaying treatment will not result in harm. Subjects who sense they are not improving may drop out of trials because they attribute lack of effect to having been treated with placebo, complicating the analysis of the study. With care, however, drop-out for lack of effectiveness can sometimes be used as a study endpoint. Although this may provide some information on drug effectiveness, such information is less precise than actual information on clinical status in subjects receiving their assigned treatment.

2.1.7.3 *Generalizability.* It is sometimes argued that any controlled trial, but especially a placebo-controlled trial, represents an artificial environment that gives results different from true "real world" effectiveness. If study populations are unrepresentative in placebo-controlled trials because of ethical or practical concerns, questions about the generalizability of study results can arise. For example, patients with more serious disease may be excluded by protocol, investigator, or patient choice from placebo-controlled trials. In some cases, only a limited member of patients or centers may be willing to participate in studies. Whether these concerns actually (as opposed to theoretically) limit generalizability has not been established.

2.1.7.4 *No comparative information.* Placebo-controlled trials lacking an active control give little useful information about comparative effectiveness, information that is of interest and importance in many circumstances. Such information cannot reliably be obtained from cross-study comparisons, as the conditions of the studies may have been quite different.

2.2 No-Treatment Concurrent Control (See Section 1.3.2)

The randomized no-treatment control is similar in its general properties and its advantages and disadvantages to the placebo-controlled trial. Unlike the placebo-controlled trial, however, it cannot be fully blinded, and this can affect all aspects of the trial, including subject retention, patient management, and all aspects of observation (see section 1.2.2). This design is appropriate in circumstances where a placebo-controlled trial would be performed, except that blinding is not feasible because the treatments themselves are so different, e.g. radiation therapy versus surgery, or because the treatment side effects are so different. When this design is used, it is desirable to have critical decisions, such as eligibility and endpoint determination or changes in management, made by an observer blinded to treatment assignment. Decisions related to data analysis, such as inclusion of patients in analysis sets, should also be made by individuals without access to treatment assignment (See ICH E9 for further discussion).

2.3 Dose-Response Concurrent Control (See Section 1.3.3)

2.3.1 Description

A dose-response study is one in which subjects are randomly assigned to one of several dosing groups, with or without a placebo group. Dose-response studies are carried out to establish the relation between dose and efficacy/adverse effects and/or to demonstrate efficacy. The first use is considered in ICH E4; the latter is the subject of this guidance. Evidence of efficacy could be based on significant differences in pair-wise comparisons between dosing groups or between dosing groups and placebo, or on evidence of a significant positive trend with increasing dose, even if no two groups are significantly different. In the latter case, however, further study may be needed to assess the effectiveness of the low doses. As noted in ICH E9, the particular approach for the primary efficacy analysis should be prespecified.

There are several advantages to inclusion of a placebo (zero-dose) group in a dose-response study. First, it avoids studies that are uninterpretable because all doses produce similar effects so that one cannot assess whether all doses are equally effective or equally ineffective. Second, the placebo group permits an estimate of absolute size of effect, although the estimate may not be very precise if the dosing groups are relatively small. Third, as the drug-placebo difference is generally larger than inter-dose differences, use of placebo may permit smaller sample sizes. The size of various dose groups need not be identical; e.g., larger samples could be used to give more precise information about the effect of smaller doses or be used to increase the power of the study to show a clear effect of what is expected to be the optimal dose. Dose-response studies can include one or more doses of an active-control agent. Randomized withdrawal designs can also assign subjects to multiple dosage levels.

2.3.2 Ability to Minimize Bias

If the dose-response study is blinded, it shares with other blinded designs an ability to minimize subject and investigator bias. When a drug has pharmacologic effects that could break the blind for some patients or investigators, it may be easier to preserve blinding in a dose-response study than in a placebo-controlled trial. Masking treatments may necessitate multiple dummies or preparation of several different doses that look alike.

2.3.3 Ethical Issues

The ethical and practical concerns related to a dose-response study are similar to those affecting placebo-controlled trials. Where there is therapy known to be effective in preventing death or irreversible morbidity, it is no more ethically acceptable to randomize deliberately to subeffective therapy than it is to randomize to placebo. Where therapy is directed at less serious conditions or where the toxicity of the therapy is substantial relative to its benefits, dose-response studies that use low, potentially subeffective doses or placebo may be acceptable to patients and investigators.

2.3.4 Usefulness of Dose-Response Studies and Quality/Validity of Inference in Particular Situations

In general, a blinded dose-response study is useful for the determination of efficacy and safety in situations where a placebo-controlled trial would be useful and has similar credibility (see section 2.1.4).

2.3.5 Modifications of Design and Combinations With Other Controls That Can Resolve Ethical, Practical, or Inferential Problems

In general, the sorts of modification made to placebo-controlled studies to mitigate ethical, practical, or inferential problems are also applicable to dose-response studies (see section 2.1.5).

2.3.6 Advantages of Dose-response Trials, Other Than Those Related to Any Difference-Showing Study

2.3.6.1 *Efficiency.* Although a comparison of a large, fully effective dose to placebo is maximally efficient for showing efficacy, this design may produce unacceptable toxicity and gives no dose-response information. When the dose-response is monotonic, the dose-response trial is reasonably efficient in showing efficacy and also yields dose-response information. If the optimally effective dose is not known, it may be more prudent to study a range of doses than to choose a single dose that may prove to be suboptimal or toxic.

2.3.6.2 *Possible ethical advantage.* In some cases, notably those in which there is likely to be dose-related efficacy and dose-related important toxicity, the dose-response study may represent a difference-showing trial that can be ethically or practically conducted even where a placebo-controlled trial could not be, because there is reason for patients and investigators to accept lesser effectiveness in return for greater safety.

2.3.7 Disadvantages of Dose-Response Study

A potential problem that needs to be recognized is that a positive dose-response trend (i.e., a significant correlation between the dose and the efficacy outcome), without significant pair-wise differences, can establish efficacy, but may leave uncertainty as to which doses (other than the largest) are actually effective. But, of course, a single-dose study poses a similar problem with respect to doses below the one studied, giving no information at all about such doses.

It should also be appreciated that it is not uncommon to show no difference between doses in a dose-response study; if there is no placebo group to provide a clear demonstration of an effect, this is a very costly "no test" outcome.

If the therapeutic range is not known at all, the design may be inefficient, as many patients may be assigned to sub-therapeutic or supratherapeutic doses.

Dose-response designs may be less efficient than placebo-controlled titration designs for showing the presence of a drug effect; they do, however, in most cases provide better dose-response information (see ICH E4).

2.4 Active Control

2.4.1 Description (See Section 1.3.4)

An active-control (positive-control) trial is one in which an investigational drug is compared with a known active drug. Such trials are usually randomized and usually double-blind. The most crucial design question is whether the trial is intended to show a difference between the two drugs or to show noninferiority/equivalence. A sponsor intending to demonstrate effectiveness by means of a trial showing noninferiority of the test drug to a standard agent needs to address the issue of the sensitivity-to-drug-effects and assay sensitivity of the trial, as discussed in section 1.5. In a noninferiority/equivalence trial, the active-control agent needs to be of established efficacy at the dose used and under the conditions of the study (see ICH E9: Statistical Principles for Clinical Trials). In general, this means it should be an agent acceptable in the region to which the studies will be submitted for the same indication at the dose being studied. A superiority study favoring the test drug, on the other hand, is readily interpretable as evidence of efficacy, even if the dose of active control is too low or the active control is of uncertain benefit (but not if it could be harmful). Such a result, however--superiority in the trial of the test agent to the control--is interpretable as actual superiority of the test drug to the control treatment only when the active control is used in appropriate patients at an optimal dose and schedule (see section 1.4.2). Lack of appropriate use of the control drug would also make the study unusable as a noninferiority study if superiority of the test drug is not shown, because assay sensitivity of the study would not be ensured (see section 1.5.4).

2.4.2 Ability to Minimize Bias

A randomized and blinded active-control trial generally minimizes subject and investigator bias, but a note of caution is warranted. In a noninferiority trial, investigators and subjects know that all subjects are getting active drug, although they do not know which one. This could lead to a biased interpretation of results in the form of a tendency toward categorizing borderline cases as successes in partially subjective evaluations, e.g., in an antidepressant study. Such biases may decrease variance and/or treatment differences and thus can increase the likelihood of an incorrect finding of equivalence.

2.4.3 Ethical Issues

Active-control trials are generally considered to pose fewer ethical and practical problems than placebo-controlled trials because all subjects receive active treatment. It should be appreciated, however, that subjects getting a new agent are not getting standard therapy (just as a placebo group is not) and may be receiving an ineffective or harmful drug. This is an important matter if the active-control therapy is known to improve survival or decrease the occurrence of irreversible morbidity. There should therefore be a sound rationale for the investigational agent. If there is not strong reason to expect the new drug to be at least as good as the standard, an add-on study (see section 2.1.5.2.1) may be more appropriate, if the conditions allow such a design.

Using a very low dose, either of the active control or of the test drug, may provide a de facto placebo that can be shown inferior to the full dose of the test drug. This, however, is only considered ethical where a placebo would also be ethical, unless there is a legitimate reason to study such low doses.

2.4.4 Usefulness of Active-Control Trials and Quality/Validity of Inference in Particular Situations

When a new drug shows an advantage over an active control, the study has inferential properties regarding the presence of efficacy equivalent to any other difference-showing trial, assuming that the active control is not actually harmful. When an active-control trial is used to show noninferiority/equivalence, there is the special consideration of sensitivity-to-drug-effects and assay sensitivity, which are considered above in section 1.5. If assay sensitivity is established, either historically (by reference to past experience

with the control drug) or by including a placebo control as well as active control, the active-control trial can assess comparative efficacy.

2.4.5 Modifications of Design and Combinations With Other Controls That Can Resolve Ethical, Practical, or Inferential Issues

As discussed earlier (section 2.1.5), active-control studies can include a placebo group, multiple-dose groups of the test drug, and/or other dose groups of the active control. Comparative dose-response studies, in which there are several doses of both test and active control, are typical in analgesic trials. The doses in active-control trials can be fixed or titrated, and both cross-over and parallel designs can be used. The assay sensitivity of a noninferiority trial can sometimes be supported by a randomized placebo-controlled withdrawal phase at the end (see section 2.1.5.2.4). Active-control superiority studies in selected populations (nonresponders to other therapy) can be very useful and are generally easy to interpret (see appendix), although the results may not be generalizable.

2.4.6 Advantages of Active-Control Trials

2.4.6.1 *Ethical/practical advantages.* The active-control design, whether intended to show noninferiority/equivalence or superiority, reduces ethical concerns that arise from failure to use drugs with documented important health benefits. It also addresses patient and physician concerns about failure to use documented effective therapy. Recruitment and IRB/IEC approval may be facilitated, and it may be possible to study larger samples. There may be fewer dropouts due to lack of effectiveness.

2.4.6.2 *Information content.* Where superiority to an active treatment is shown, active-control studies are readily interpretable regarding evidence of efficacy. The larger sample sizes needed are sometimes more achievable and acceptable in active-control trials and can provide more safety information. Active-control trials also can, if properly designed, provide information about relative efficacy.

2.4.7 Disadvantages of Active-Control Trials

2.4.7.1 *Information content.* See section 1.5 for discussion of the problem of assay sensitivity and the ability of the trial to support an efficacy conclusion in noninferiority/equivalence trials. Even when

assay sensitivity is supported and the study is suitable for detecting efficacy, there is no direct assessment of absolute effect size and greater difficulty in quantitating safety outcomes as well.

2.4.7.2 *Large sample size.* Generally, in noninferiority trials, the margin of difference that needs to be excluded is chosen conservatively, first, because the smallest effect of the active control expected in trials will ordinarily be used as the estimate of its effect and, second, because there will usually be an intent to rule out loss of more than some reasonable fraction (see section 1.5.2) of the control drug effect, leading to a still smaller margin. Because of the need for conservative assumptions about control drug effect size, sample sizes may be very large. In a difference-showing active-control trial, the difference between two drugs is always smaller, often much smaller, than the expected difference between drug and placebo, again leading to large sample sizes.

2.5 External Control (Historical Control)

2.5.1 Description

An externally controlled trial is one in which the control group consists of patients who are not part of the same randomized study as the group receiving the investigational agent, i.e., there is no concurrently randomized comparative group. The control group is thus not derived from exactly the same population as the treated population. Usually, the control group is a well-documented population of patients observed at an earlier time (historical control) at another institution, or even at the same institution but outside the study. An external-control study could be a superiority study or an equivalence study. Sometimes certain patients from a larger experience are selected as a control group on the basis of particular characteristics that make them similar to the treatment group; there may even be an attempt to ''match'' particular control and treated patients.

So-called ''baseline-controlled studies'' are a variety of externally controlled trials; these are sometimes thought to use ''the patient as his own control,'' but that is logically incorrect. In fact, the comparator group is an estimate of what would have happened in the absence of therapy to the patients. Both baseline-controlled trials and studies that use a more complicated on-off-on (cross-over) design, but that do not include a concurrently randomized control group, are of this type. As noted, in these studies

the observed changes from baseline or between study periods are always compared, at least implicitly, to some estimate of what would have happened without the intervention. Such estimates are generally made on the basis of "general knowledge," without reference to a specific control population. Although in some cases this is plainly reasonable, e.g., when the effect is dramatic, occurs rapidly following treatment, and is unlikely to have occurred spontaneously (e.g., general anesthesia, cardioversion, measurable tumor shrinkage), in most cases it is not so obvious and a specific historical experience should be sought. Designers and analysts of such trials need to be aware of the risks of this type of control and should be prepared to support its use.

2.5.2 Ability to Minimize Bias

Inability to control bias is the major and well-recognized limitation of externally controlled trials and is sufficient in many cases to make the design unsuitable. It is always difficult, in many cases impossible, to establish comparability of the treatment and control groups and thus to fulfill the major purpose of a control group (see section 1.2). The groups can be dissimilar with respect to a wide range of factors, other than the study drug, that could affect outcome, including demographic characteristics, diagnostic criteria, stage or duration of disease, concomitant treatments, and observational conditions (such as methods of assessing outcome, investigator expectations). Blinding and randomization are not available to minimize bias when external controls are used. It is well documented that untreated historical-control groups tend to have worse outcomes than an apparently similar control group in a randomized study, primarily because of selection bias. Control groups in a randomized study should meet certain criteria to be entered into the study, criteria that are generally more stringent and identify a less sick population than is typical of external-control groups. The group is often identified retrospectively, leading to potential bias in its selection. A consequence of the recognized inability to control bias is that the persuasiveness of findings from externally controlled trials depends on obtaining much more extreme levels of statistical significance and much larger estimated differences between treatments than would be considered persuasive in concurrently controlled trials.

The inability to control bias restricts use of the external-control design to situations in which the effect of treatment is dramatic and the usual course of the disease highly predictable. In addition, use of external controls should be limited to cases in which the endpoints are objective and the impact of baseline and treatment variables on the endpoint is well characterized.

As noted, the lack of randomization and blinding, and the resultant problems with lack of assurance of comparability of test group and control group, make the likelihood of substantial bias inherent in this design and impossible to quantitate. Nonetheless, some approaches to design and conduct of externally controlled trials could lead them to be more persuasive and potentially less biased. A control group should be chosen for which there is detailed information, including, where needed, individual patient data regarding demographics, baseline status, concomitant therapy, and course on study. The control patients should be as similar as possible to the population expected to receive the test drug in the study and should have been treated in a similar setting and in a similar manner, except with respect to the study therapy. Study observations should utilize timing and methodology similar to those used in the control patients. To reduce selection bias, selection of the control group should be made before performing comparative analyses; this may not always be feasible, as outcomes from these control groups may have been published. Any matching on selection criteria or adjustments made to account for population differences should be specified prior to selection of the control and performance of the study. Where no obvious single "optimal" external control exists, it may be advisable to study multiple external controls, providing that the analytic plan specifies conservatively how each will be utilized in drawing inferences (e.g., study group should be substantially superior to the most favorable control to conclude efficacy). In some cases, it may be useful to have an independent set of reviewers reassess endpoints in the control group and in the test group in a blinded manner according to common criteria.

2.5.3 Ethical Issues

When a drug is intended to treat a serious illness for which there is no satisfactory treatment, especially if the new drug is seen as promising on the basis of theoretical considerations, animal data, or early human experience, there may be understandable reluctance to perform a comparative study with a concurrent

control group of patients who would not receive the new treatment. At the same time, it is not responsible or ethical to carry out studies that have no realistic chance of credibly showing the efficacy of the treatment. It should be appreciated that many promising therapies have had less dramatic effects than expected or have shown no efficacy at all when tested in controlled trials. Investigators may, in these situations, be faced with very difficult judgments. It may be tempting in exceptional cases to initiate an externally controlled trial, hoping for a convincingly dramatic effect, with a prompt switch to randomized trials if this does not materialize.

Alternatively, and generally preferably, in dealing with serious illnesses for which there is no satisfactory treatment, but where the course of the disease cannot be reliably predicted, even the earliest studies should be randomized. This is usually possible when studies are carried out before there is an impression that the therapy is effective. Studies can be monitored by independent data monitoring committees so that dramatic benefit can be detected early. Despite the use of a single-treatment group in an externally controlled trial, a placebo-controlled trial is usually a more efficient design (needing fewer subjects) in such cases, as the estimate of control group outcome generally needs to be made conservatively, causing need for a larger sample size. Great caution (e.g., applying a more stringent significance level) is called for because there are likely to be both identified and unidentified or unmeasurable differences between the treatment and control groups, often favoring treatment. The concurrently controlled trial can detect extreme effects very rapidly and, in addition, can detect modest, but still valuable, effects that would not be credibly demonstrated by an externally controlled trial.

2.5.4 Usefulness of Externally Controlled Trials and Quality/Validity of Inference in Particular Situations

An externally controlled trial should generally be considered only when prior belief in the superiority of the test therapy to all available alternatives is so strong that alternative designs appear unacceptable and the disease or condition to be treated has a well-documented, highly predictable course. It is often possible, even in these cases, to utilize alternative, randomized, concurrently controlled designs (see section 2.1.5 and appendix).

Externally controlled trials are most likely to be persuasive when the study endpoint is objective, when the outcome on treatment is markedly different from that of the external control and a high level of statistical significance for the treatment-control comparison is attained, when the covariates influencing outcome of the disease are well characterized, and when the control closely resembles the study group in all known relevant baseline, treatment (other than study drug), and observational variables. Even in such cases, however, there are documented examples of erroneous conclusions arising from such trials.

When an external-control trial is considered, appropriate attention to design and conduct may help reduce bias (see section 2.5.2).

2.5.5 Modifications of Design and Combinations With Other Controls That Can Resolve Ethical, Practical or Inferential Problems

The external-control design can incorporate elements of randomization and blinding through use of a randomized placebo-controlled withdrawal phase, often with early-escape provisions, as described earlier (see section 2.1.5.2.4). The results of the initial period of treatment, in which subjects who appear to respond are identified and maintained on therapy, are thus "validated" by a rigorous, largely assumption- and bias-free study.

2.5.6 Advantages of Externally Controlled Trials

The main advantage of an externally controlled trial is that all patients can receive a promising drug, making the study more attractive to patients and physicians.

The design has some potential efficiencies (smaller sample size) because all patients are exposed to test drug, of particular importance in rare diseases.

2.5.7 Disadvantages of Externally Controlled Trials

The externally controlled study cannot be blinded and is subject to patient, observer, and analyst bias, major disadvantages. It is possible to mitigate these problems to a degree, but even the steps suggested in section 2.5.2 cannot resolve such problems fully, as treatment assignment is not randomized and comparability of control and treatment groups at the start of treatment, and comparability of treatment

of patients during the trial, cannot be ensured or well assessed. It is well documented that externally controlled trials tend to overestimate efficacy of test therapies.

3.0 Choosing the Control Group

Figure 1 and Table 1 provide a decision tree for choosing among different types of control groups. Although the table and figure focus on the choice of control to demonstrate efficacy, some designs also allow comparisons of test and control agents. The choice of control can be affected by the availability of therapies and by medical practices in specific regions.The potential usefulness of the principal types of control (placebo, active, and dose-response) in specific situations and for specific purposes is shown in Table 1. The table should be used with the text describing the details of specific circumstances in which potential usefulness can be realized. In all cases, it is presumed that studies are appropriately designed. External controls are so distinct a case that they are not included in the table. In the table, a P notation refers to the need to make a convincing case that the study has assay sensitivity.

In general, evidence of efficacy is most convincingly demonstrated by showing superiority to a concurrent control treatment. If a superiority trial is not feasible or is inappropriate for ethical or practical reasons, and if a defined treatment effect of the active control is regularly seen (e.g., as it is for antibiotics in most situations), a noninferiority/equivalence study can be utilized and can be persuasive. Use of this design calls for close attention to the issue of sensitivity to drug effects in active-control noninferiority trials of the condition being studied and to the assay sensitivity of the particular study carried out (see section 1.5).

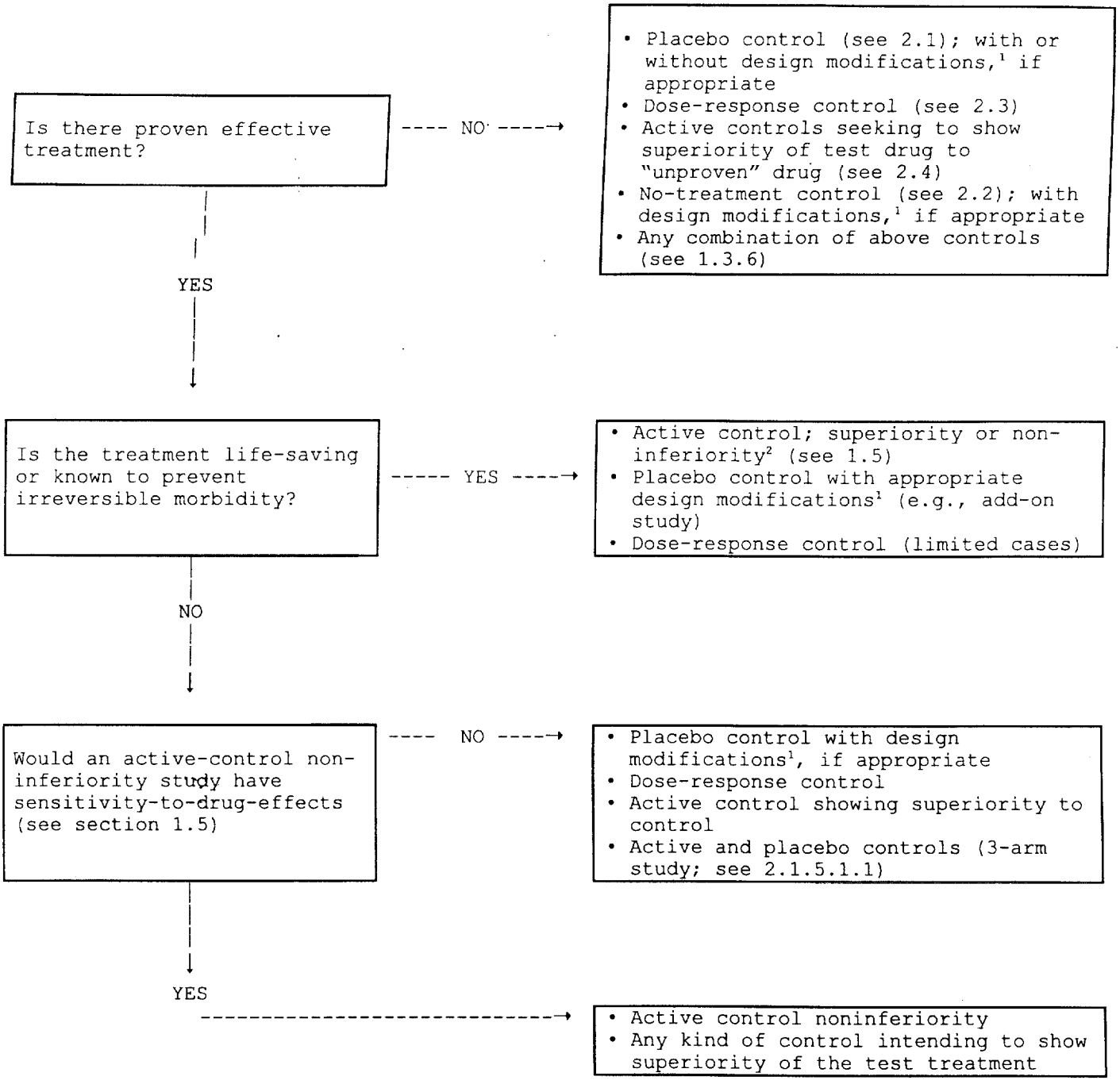
Table 1.--Usefulness of Specific Control Types in Various Situations

	Type of Control							
Trial Objective	Placebo	Active non-inferiority	Active Superiority	Dose Response (D/R)	Placebo + Active	Placebo + D/R	Active + D/R	Placebo + Active + D/R
Measure absolute effect size	Y	N	N	N	Y	Y	N	Y
Show existence of effect	Y	P	Y	Y	Y	Y	Y	Y
Show Dose Response relationship	N	N	N	Y	N	Y	Y	Y
Compare therapy (show superiority, show non-inferiority)	N	P	Y	N	Y	N	P	Y
Show assay sensitivity[1]	Y	N	Y	Y	Y	Y	Y	Y

Y=Yes, N=No, P=Possible, depending on a showing that this type of trial will have sensitivity-to-drug-effect.
[1] Through the direct demonstration within the trial of the ability to demonstrate differences between treatments.

Figure 1 Choosing the Concurrent Control for Demonstrating Efficacy

This figure shows the basic logic for choosing the control group; the decision may depend on the available drugs or medical practices in the specific region.



[1] Add-on, replacement, early escape, brief placebo period, and randomized withdrawal (see section 2.1.5.2).
[2] If known sensitivity-to-drug-effects (see section 1.5).

APPENDIX

Studies of Efficacy in Subsets of the Whole Population; Enrichment

1.0 Introduction

Ideally, the effect of a drug should be known in general and in relevant demographic and other subsets of the population, such as those defined by disease severity or other disease characteristics. To the extent study patients are not a random sample of the patients who will be treated with the drug once it is marketed, the generalizability of the results can be questioned. Even if the overall result is obtained in a representative sample, however, that does not suggest the result is the same in all people. If subject selection criteria can identify people more likely to respond to therapy (e.g., high renin hypertensives to beta blockers), we consider therapy more rational and the drug more useful.

Subjects entering clinical studies are in fact almost never a random sample of the potential treatment population, and they are not treated exactly as a nonstudy patient would be treated. They must give informed consent, be able to follow instructions, and be able to get to the clinic. They are sometimes assessed for likelihood of complying with treatment. They are usually not very debilitated and generally are without complicated or life-threatening illness, unless those conditions are being studied. They are usually selected using particularly stringent diagnostic criteria that make it very certain they actually have the disease to be treated (more likely than in clinical practice). Lead-in periods are often used to exclude subjects who improve spontaneously or whose relevant functional measures (blood pressure, exercise tolerance) are too variable. Of course, the entire setting of trials is artificial in varying degrees, generally directed toward reducing unwanted variability and increasing study efficiency.

All of these departures from a truly unselected population of people likely to receive the drug are directed at identifying and including subjects likely to make a "good assay population." They can be considered methods of "enrichment" of the population, modifications of a truly random sample of potential users to produce a population of subjects more likely to discriminate between an active and an inactive therapy. The kinds of enrichment described above are widely accepted and "benign," i.e., it seems likely that results in such a population will be of general applicability, at least to patients with good compliance.

There is a view, however, that in-use ''effectiveness'' may often be different from the artificial ''efficacy'' established in these enriched ''efficacy'' trials.

There are other kinds of enrichment that could also be useful but that would more clearly alter the inference that could be drawn from the results. This should not discourage their use but should encourage attention to what such studies do, and do not, show. Some enrichments of potential value include:

1.1 Studies of Patients Nonresponsive to, or Intolerant of, Other Therapy

In this kind of study, patients failing therapy on a drug, or failing to tolerate it acceptably, are randomized to the failed or poorly tolerated therapy or to the investigational treatment. Greater efficacy (or better tolerance) of the new therapy shows that the drug is useful in failures on the other therapy. This is a valuable showing if, e.g., the drug is relatively toxic and intended for a ''second-line'' use, but it does not show that the new therapy is superior in general, and such studies need to be carefully interpreted. By selecting study patients who will only infrequently respond to the control agent or who are very likely to have a particular adverse effect of the control drug, the design facilitates showing the second drug's advantage in that circumstance. A direct comparison of the two drugs in an unselected population that could contain responders to both drugs would need to be much larger to show a difference between the treatments, even if there was an overall advantage of the new drug. Moreover, it could be that each drug has a similar rate of nonresponders (but the other drug works in some of these), so that no difference could be seen in a direct comparison in unselected subjects.

In this design, it is usually critical to randomize the nonresponders or intolerants to both the new agent and the failed agent, rather than simply place the failures on the new drug. Patients who failed previously may ''respond'' to the failed drug when it is readministered in a clinical trial, or may tolerate the previously poorly tolerated drug in the new circumstance. This can present a problem. In the ''intolerance'' case, although subjects can be randomized to a drug that has caused certain kinds of intolerance, they cannot be randomized to a drug that would endanger them if administered (e.g., if the intolerance was anaphylaxis, liver necrosis). Similarly, in the nonresponder case, patients cannot be

restudied on the failed drug if failure would lead to harm. In some cases, the prior experience may be an adequate control (e.g., failure of a tumor to respond), a baseline-controlled study design.

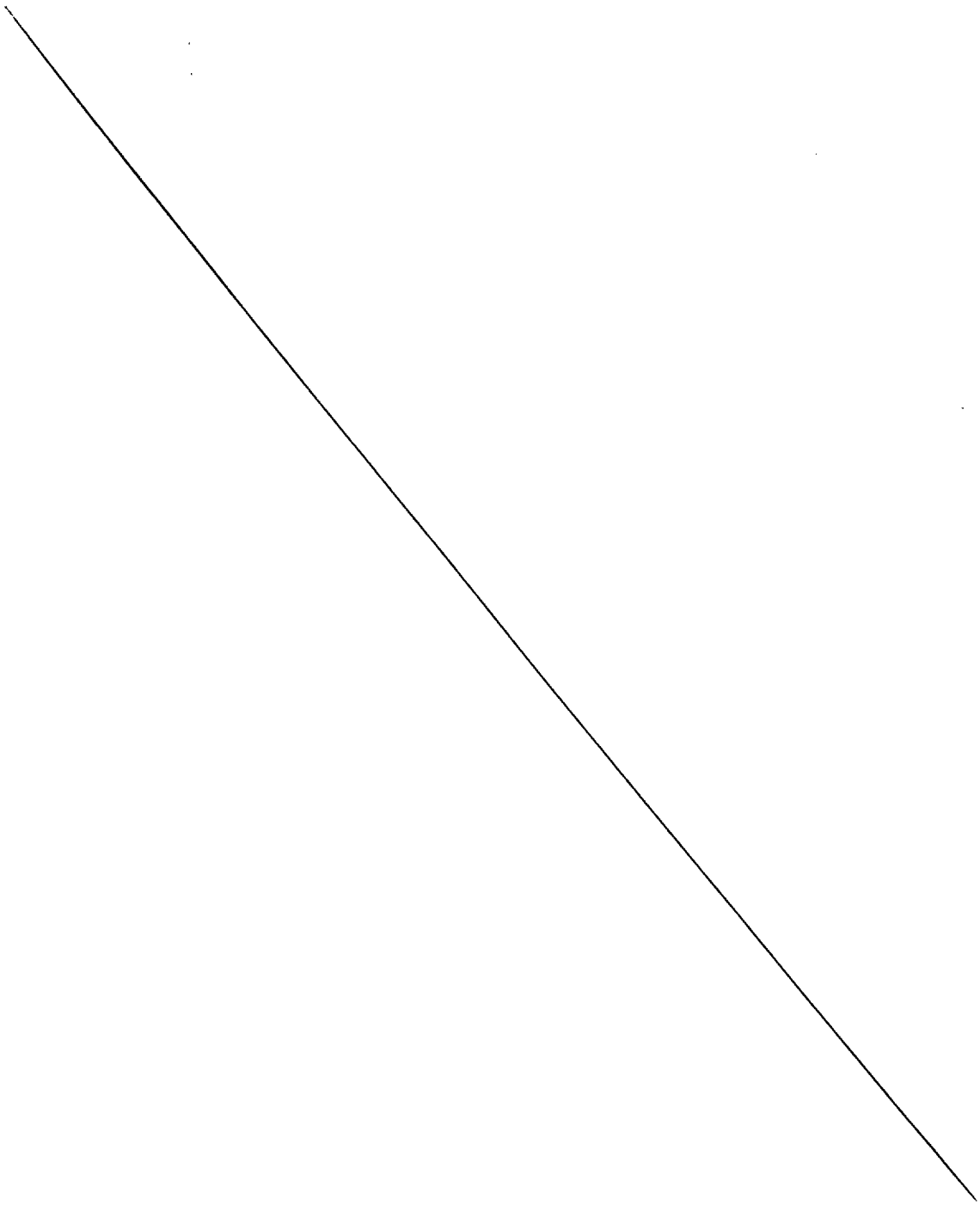
1.2 Studies in Likely or Known Responders

If patients cannot respond to the main pharmacologic effect of the drug, they cannot be expected to show a clinical response. Thus, subjects with no blood pressure response to sublingual nitroglycerin have been excluded from trials of organic nitrates, as they show no ability to respond to the mechanism of action of these drugs and including them would only dilute the drug effect. A similar approach was used in Cardiac Arrhythmia Suppression Trial (CAST). Only subjects responding to encainide or flecainide with a 70 percent reduction in ventricular premature beats (VPB's) were randomized to the mortality phase of the study because there was no reason to include people who could not possibly benefit (i.e., people with no VPB reduction). It is important in such cases to record the number of subjects screened in order to construct the study population so that users of the drug will have a reasonable expectation of what they will encounter. It will often be appropriate to incorporate similar selection criteria in labeling the drug for use.

The nitroglycerin and CAST enrichment approaches were generally accepted. A potentially more controversial enrichment procedure would be to identify responders in an initial open phase, withdraw treatment, then carry out a randomized study in the responders. This could be a useful approach when efficacy has proved difficult to demonstrate. For example, it has been difficult to obtain evidence that gut motility-modifying agents are effective in gastroesophageal reflux disease, perhaps because there are unrecognized pathophysiologic subsets of patients, some of which can respond and some of which cannot. It seems possible that identifying apparent responders clinically, then randomizing the apparent responders to drug and placebo treatments, would best utilize both clinical observation and rigorous design.

In seeking dose-response information, little is to be learned from studying the drug in a population of nonresponders (although one would want to know the proportion of the population that is nonreponsive). Such studies might better be carried out in known responders to the drug. Similarly, in evaluating a drug

of a particular class, studies including only known responders to the class might be more likely to detect an effect of the drug or to show differences between members of the class.



Finally, it should be appreciated that randomized withdrawal studies (see section 2.1.5.2.4), and studies of maintenance treatment in general, are often studies in known responders and can therefore be expected to show greater effect than studies in an unselected population.

Dated: 9/16/99

September 16, 1999

Margaret M. Dotzel
Acting Associate Commissioner for Policy

[FR Doc. 99–???? Filed ??–??–99; 8:45 am]

BILLING CODE 4160–01–F